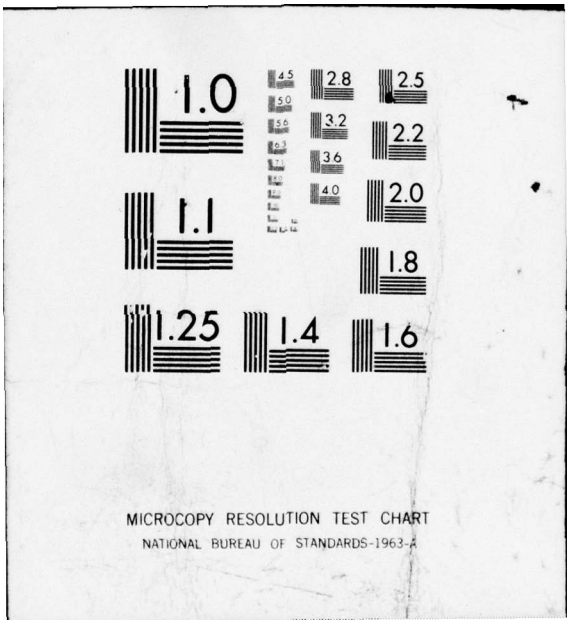
AD-A031 676 RAND CORP SANTA MONICA CALIF F/G 8/3
ESTIMATION OF BOUNDARY CONDITIONS FOR COASTAL MODELS,(U)
SEP 74 S K LIU, J J LEENDERTSE, J VOOGT

UNCLASSIFIED P-5299 NL

| OF |
AD
A031676

END

DATE
FILMED
12-76


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(2) FC

(6) ESTIMATION OF BOUNDARY CONDITIONS FOR COASTAL MODELS

(10) S. K./Liu
J. J./Leendertse
J./Voogt

(11) September 1974

(12) 20p.

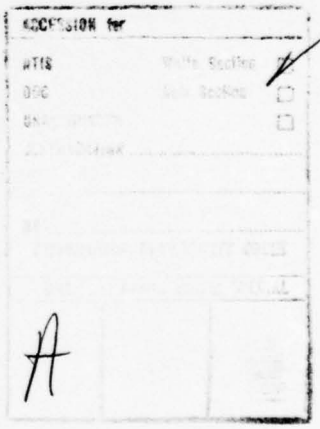
(14) P-5299

296 600 AB

The Rand Corporation
Santa Monica, California 90406

ESTIMATION OF BOUNDARY CONDITIONS FOR COASTAL MODELS

S. K. Liu[1]
J. J. Leendertse[1]
J. Voogt[2]



This paper was presented at the 14th International Conference on Coastal Engineering, Copenhagen, June 24-28, 1974.

---

[1]The Rand Corporation, Santa Monica, California 90406 U.S.A.

[2]Rijkswaterstaat, Dienst Informatieverwerking, The Hague, The Netherlands.

ESTIMATION OF BOUNDARY CONDITIONS FOR COASTAL MODELS

by

S. K. Liu,[1] J. J. Leendertse,[1] and J. Voogt[2]

ABSTRACT

In this study, frequency response and transfer function techniques are used together with cross-spectral and fast Fourier transform methods to determine the proper boundary values for computing the flow field of a coastal sea. Tide data containing considerable perturbations from swell and meteorological disturbances are analyzed.

In computing the frequency response estimates, the effect of noise in the input is treated by a cancelling technique and by the choice of a reference station to evaluate the interdependencies among the other stations at the boundary. The usefulness of the network frequency response function is threefold: (1) future conditions can be simulated using observed water levels at any single location, (2) boundary information for models of different grid size can be obtained by interpolation, and (3) missing data at a given location can be estimated optimally using data at neighboring stations and the network response function. The paper discusses an example of such an application, the determination of a boundary of a two-dimensional model of Jamaica Bay, New York City, U.S.A.

INTRODUCTION

One of the major difficulties in coastal and estuarine hydrodynamic computation is obtaining good boundary information. For example, the computation requires the time histories of water levels at open boundaries as one of the major input forcing functions from which is derived the internal flow field. Field measurements at boundaries, as well as within the prototype, are also needed during various phases of model development and adjustment. However, such field data often contain noise generated by instruments, meteorological disturbances, or short-period waves. Often, part of the records of the critical period may even be missing. This paper deals mainly with problems such as these encountered frequently in hydrodynamic and water quality modeling. The analyses used are the estimation of network frequency response function, cross-spectral computation, noise cancellation, and numerical convolution.

---

[1]The Rand Corporation, Santa Monica, California 90406 U.S.A.

[2]Rijkswaterstaat, Dienst Informatieverwerking, The Hague, The Netherlands.

ESTIMATION OF NETWORK FREQUENCY RESPONSE FUNCTIONS

The frequency response function H(f), as its name indicates, describes the amplitude and phase relationship between one fluctuation with a certain frequency f as input and the resulting fluctuation as output.

In determining network frequency response relationships, the statistical approach using cross-spectral estimates is used instead of the classic method employing the deterministic Fourier or Laplace transforms, because the sampling variability, confidence limits, and the phase of the frequency response function can only be estimated using cross-spectral procedures.

A description of the computational method used in this paper is presented by Liu,[1] and by Leendertse and Liu.[2,3] Reference is also made to the handbook of Jenkins and Watts[4] and the thesis of Goodman[5] on this subject. A brief outline of the computational method is given below.

With respect to the interdependency (or the lack of it) between two random time series, x_t and y_t, we see that if the random processes are jointly stationary such that the joint distribution depends only on time differences, then the degree of interdependency can be measured by the cross-covariance function. In discrete time this is defined as

$$\hat{\gamma}_{xy}(k) = \frac{1}{n-k} \sum_{t=1}^{n-k} (x_t - \bar{x})\,(y_{t+k} - \bar{y})$$

$$\hat{\gamma}_{yx}(k) = \frac{1}{n-k} \sum_{t=1}^{n-k} (y_t - \bar{y})\,(x_{t-k} - \bar{x}) \qquad (1)$$

for k = 0, 1, 2, ... m, where m is the largest time lag chosen, n the total number of data points, and $\bar{x},\bar{y}$ the mean values of the series {x},{y}.

The relationship between these two stochastic processes can also be expressed by the integral equation

$$y(t) - \mu = \int_0^\infty h(\tau)[X(t-\tau) - \mu_x]\,d\tau + N(t) \qquad (2)$$

where $h(\tau)$ is the impulse response function, μ_x, μ_y are the mean values of the two processes, and N(t) is the uncorrelated error term.

Wiener[6] showed that the optimal estimates of $h(\tau)$ should satisfy the following integral equation:

$$\gamma_{xy}(t) = \int_0^\infty \hat{h}(\tau)\,\gamma_{xx}(t-\tau)\,d\tau \qquad (3)$$

The solution to Eq. (3) may be obtained by Fourier transformation. Because covariance function and spectral density function form a Fourier transform pair, thus

$$P_{xy}(f) = H(f) \cdot P_{xx}(f) \quad \text{or} \quad H(f) = P_{xy}(f)/P_{xx}(f) \tag{4}$$

where $P_{xy}(f)$ is the cross-spectrum between x(t) and y(t), $P_{xx}(f)$ is the auto-spectrum of x(t), and the complex valued function $H(f)$ is the frequency response function. The possibility of using a statistical approach using spectral densities to give the frequency response was suggested by Lee.[7] However, it was later[5] that a quantitative basis for applying the method with finite sample records corrupted by measurement noise became available.

For computing the cross-spectral density function estimate by a numerical Fourier transform, the even and odd parts of the cross-covariance function are determined by

$$\hat{A}(k) = \tfrac{1}{2}[\gamma_{xy}(k) + \gamma_{yx}(k)] \tag{5}$$

$$\hat{B}(k) = \tfrac{1}{2}[\gamma_{xy}(k) - \gamma_{yx}(k)] \tag{6}$$

from which the co-spectral density function is estimated:

$$\hat{C}_{xy}(f) = 2\tau\left[\hat{A}(o) + 2\sum_{k=1}^{k=m-1} \hat{A}(k) \cos(2\pi fk\tau) + \hat{A}(m) \cos(2\pi fm\tau)\right] \tag{7}$$

The quadrature spectral density function is estimated by

$$\hat{Q}_{xy}(f) = 2\tau\left[2\sum_{k=1}^{k=m-1} \hat{B}(k) \sin(2\pi fk\tau) + \hat{B}(m) \sin(2\pi fk\tau)\right] \tag{8}$$

The spectral density functions of input and output are determined in a similar manner. If, in Eq. (1), the output series is replaced by the input series, we obtain the auto-covariance function of the input,

$$\gamma_{xx}(k) = \frac{1}{n-k} \sum_{t=1}^{t=n-k} (x_t - \bar{x})(x_{t+k} - \bar{x}) \tag{9}$$

from which the input spectral density function is determined:

$$\hat{P}_{xx}(k) = 2\tau\left[\gamma_{xx}(0) + 2\sum_{k=1}^{k=m-1} \gamma_{xx}(k) \cos(2\pi fk\tau) + \gamma_{xx}(m) \cos(2\pi fm\tau)\right] \tag{10}$$

The output spectral density function is determined similarly.

The time interval τ used in the analysis influences the highest frequency that can be determined by the analysis method. At least two samples per cycle are required to define a frequency component in a data set; thus the highest frequency determined is

$$f_c = \frac{1}{2\tau} \tag{11}$$

This frequency is the so-called Nyquist frequency. If higher frequencies are present in the data, these are aliased as lower frequencies.

The spectral density functions are determined for particular frequencies f. These frequencies are calculated only at the special discrete frequencies of harmonic number k, where

$$f = \frac{kf_c}{m} \quad k = 0, 1, 2, 3 \ldots m \tag{12}$$

The total number of discrete frequencies determined depends on the maximum lag $m\tau$.

For each spectral function, the discrete value found is a kind of average value in a certain range or band. The bandwidth for the computations is

$$B = \frac{1}{m\tau} \tag{13}$$

One would tend to determine the spectral functions with small bandwidth -- thus in much detail -- by choosing a large value for m. Unfortunately, this considerably affects the accuracy of the result.

It should be understood that the analysis method gives estimates of the function only. Since we are dealing with data that is not deterministic, each sample record used for analysis differs from another and the results obtained from these records will also differ somewhat.

The estimates of spectral density functions described above are so-called "raw" estimates, which have certain undesirable properties. If a strong periodic component is present, the analysis may show small positive and negative values in the frequency bands adjacent to that in which the periodic component is present. This phenomenon is called "leakage," and the negative values it produces are particularly bothersome. To counter it, frequency smoothing is used, by which the estimate at a particular frequency is computed as a weighted average of the particular frequency and the adjacent frequency.

For example, the smoothed co-spectral density function can be taken as

$$\bar{\hat{C}}_{xy}(f) = \bar{\hat{C}}_{xy}\left(\frac{kf_c}{m}\right) = .25\hat{C}_{xy}\left(\frac{(k-1)f_c}{m}\right) + .5\hat{C}_{xy}\left(\frac{kf_c}{m}\right) + .25\hat{C}_{xy}\left(\frac{(k+1)f_c}{m}\right) \tag{14}$$

This frequency smoothing is called "hanning," which is equivalent to the Tukey lag window.[4] Other methods of minimizing the effect of leakage are available,[1,4] but these are not applied in this investigation.

For the absolute value of the smoothed cross-spectral density we obtain the following estimate:

$$|\bar{\hat{P}}_{xy}(f)| = \left(\bar{\hat{C}}_{xy}(f)^2 + \bar{\hat{Q}}_{xy}(f)^2\right)^{\frac{1}{2}} \tag{15}$$

Subsequently, the estimated amplitude of the frequency response function is

$$|\hat{H}(f)| = \bar{\hat{P}}_{xy}(f) / \bar{\hat{P}}_{xx}(f) = \hat{A}_{xy}(f) \tag{16}$$

and its estimated phase spectrum is

$$\hat{\phi}(f) = \tan^{-1}\left[\bar{\hat{Q}}_{xy}(f)/\bar{\hat{C}}_{xy}(f)\right] \tag{17}$$

The frequency response function thus obtained is an optimal estimate (in a least square sense), assuming that the system is linear. Even if we assume that we have a linear system whose input {x} may be measured exactly, the output may still contain measurement errors. In the case analyzed here, the output may be influenced by wind and system nonlinearities. The measured output then contains the transformed input signal plus measurement noise, etc. (in our case noise caused by wind and nonlinearities).

It is now possible to introduce a measure of the linear relation between the two series, called the coherency function $\Omega_{xy}(f)$. The squared coherency is estimated to be

$$\Omega^2_{xy}(f) = \frac{|\bar{\hat{P}}_{xy}(f)|^2}{\bar{\hat{P}}_{xx}(f)\ \bar{\hat{P}}_{yy}(f)} \tag{18}$$

If the system is completely linear, the squared coherency is unity; if the two time series are completely uncorrelated, then the coherency would be zero. If the coherency is less than unity but greater than zero, then there is noise in the measurements, the system is not linear, or the output {y} of the system is due to an input {x} as well as other inputs.

In determining the behavior of a system it is often useful to see how the noise is distributed over the frequency range. The estimated spectral density function of the noise is expressed by

$$\bar{\hat{P}}_{x\Delta y}(f) = [1 - \Omega^2_{xy}(f)]\ \bar{\hat{P}}_{yy}(f) \tag{19}$$

In a strict sense, when the cross-spectral estimates $\hat{P}_{xy}(f)$ are used to determine the frequency response function, the measurement noise at the input will cause the frequency response function to be underestimated, as can be seen from the equation

$$\hat{H}(f) = \bar{\hat{P}}_{xy}(f)/\bar{\hat{P}}_{xx}(f) = \bar{\hat{P}}_{xy}(f)/[P_{xx}(f) + P_{nnx}(f)] < H(f) \tag{20}$$

in which $P_{xx}(f)$ is the true value of input spectra and $P_{nnx}(f)$ is the spectrum of extraneous noise components in the input. Notice that the uncorrelated noise at the output does not cause bias.

The amount of measurement noise and the error due to it can be estimated by computing the amplitude function forward and backward between two sets of records (i.e., switching x series and y series in the computational sequence). By going from y to x, the noise in y would cause the transfer from y to x to be underestimated, which is equivalent to the overestimation of x in the relative amplification factor. However, by going from x to y, the noise in y no longer influences the value of backward transfer, but the noise in x is now the important factor. In each direction the transfer function is underestimated. Therefore, the best estimate of the transfer function from y to x is

$$\tfrac{1}{2}[1/A_{yx}(f) + A_{xy}(f)] \qquad (21)$$

The error would be cancelled out if the random measurement noise level in both records were about the same.

The computed coherency $\hat{\Omega}^2_{xy}(f)$ also contains some bias if the phase difference between two stations is appreciable. This bias can be reduced by a process called alignment (see Ref. 4 for details), using the peak of the cross-covariance as a guide to make the required shift.

DATA ANALYSIS

One of the applications of the network frequency response analysis is estimating open boundary conditions for a two-dimensional mathematical model of the Netherlands coast in the North Sea (Fig. 1). In order to determine the proper boundary conditions for the model, 29 bottom pressure recorders were installed by the Netherlands Rijkswaterstaat during the months of May and June 1971. Hourly water level data were first analyzed without astronomical prejudices, thus allowing for all possible frequencies and their higher harmonics that were present.

The frequency domain mapping was carried out using arbitrary-radix algorithms of fast Fourier transforms. During the transformation, Tukey's[8] interim data taper window was applied to eliminate leakage from the peaks. The Fourier line spectra for stations U (reference station) and A_1 are shown in Fig. 2. The contribution from the meteorological disturbances, located in the frequency range below 0.04 per hour, and the higher harmonic components induced by the diurnal-semidiurnal components can be noted in the graphs.

Frequency response analyses and cross-spectral computations were then carried out between the reference station U and the 21 stations located at the model open boundary. The graphic results from a typical analysis are shown in Fig. 3. In the top row the adjusted bottom pressures at station A are shown. The computed spectra at station A are also shown. It will be noted that the higher harmonics of the lunar component, which are the quarter-diurnal tide (M_4 at f = .16 hr^{-1}) and the sixth-diurnal

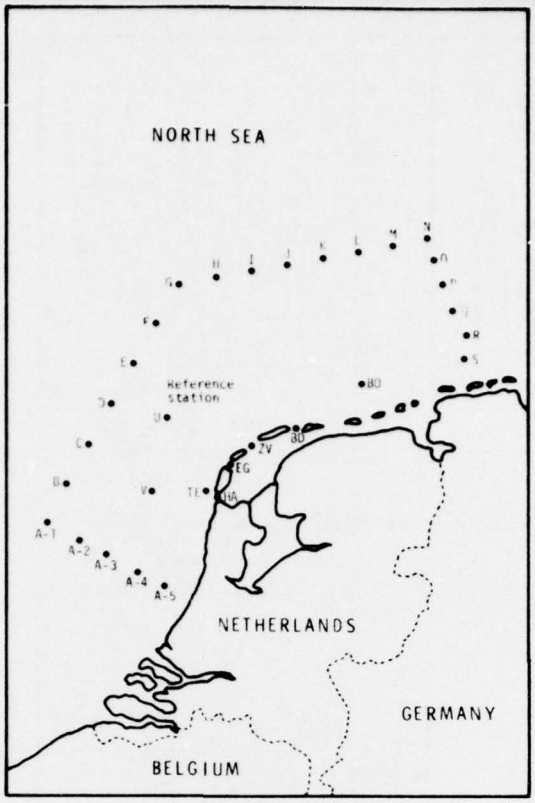


Figure 1 Relative location of tide gauges

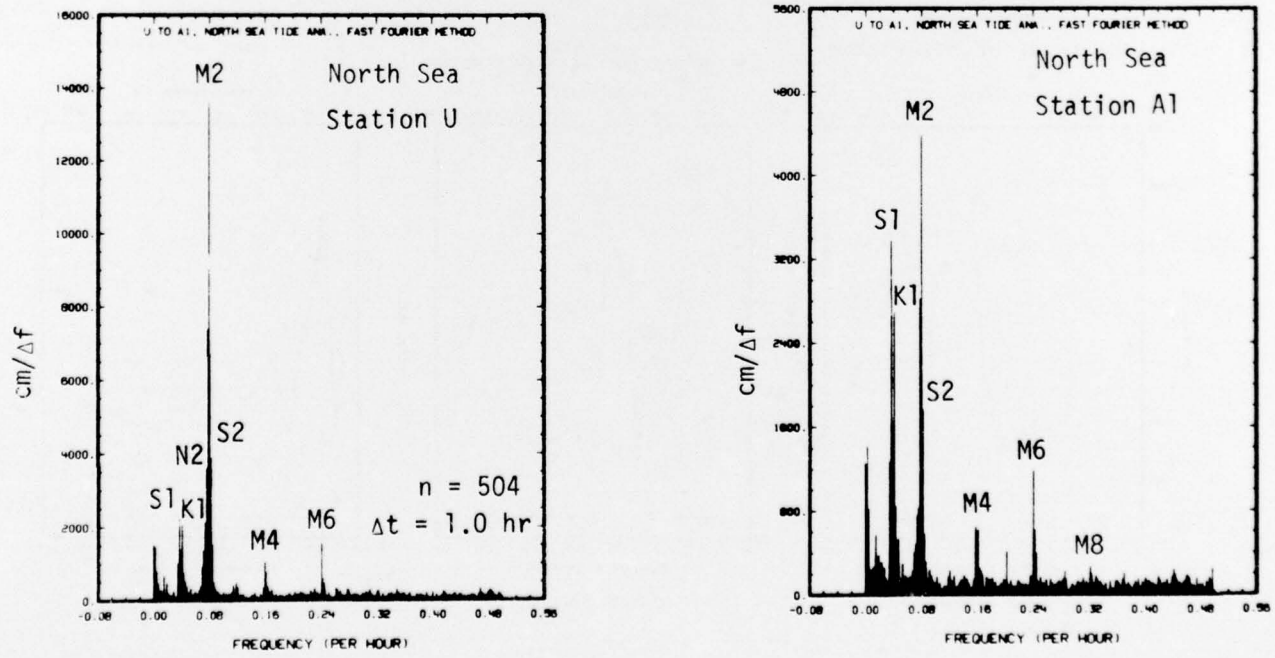


Figure 2 Analysis of tidal components and other oscillations in the observed data using fast Fourier transform and cosine data tapering filter

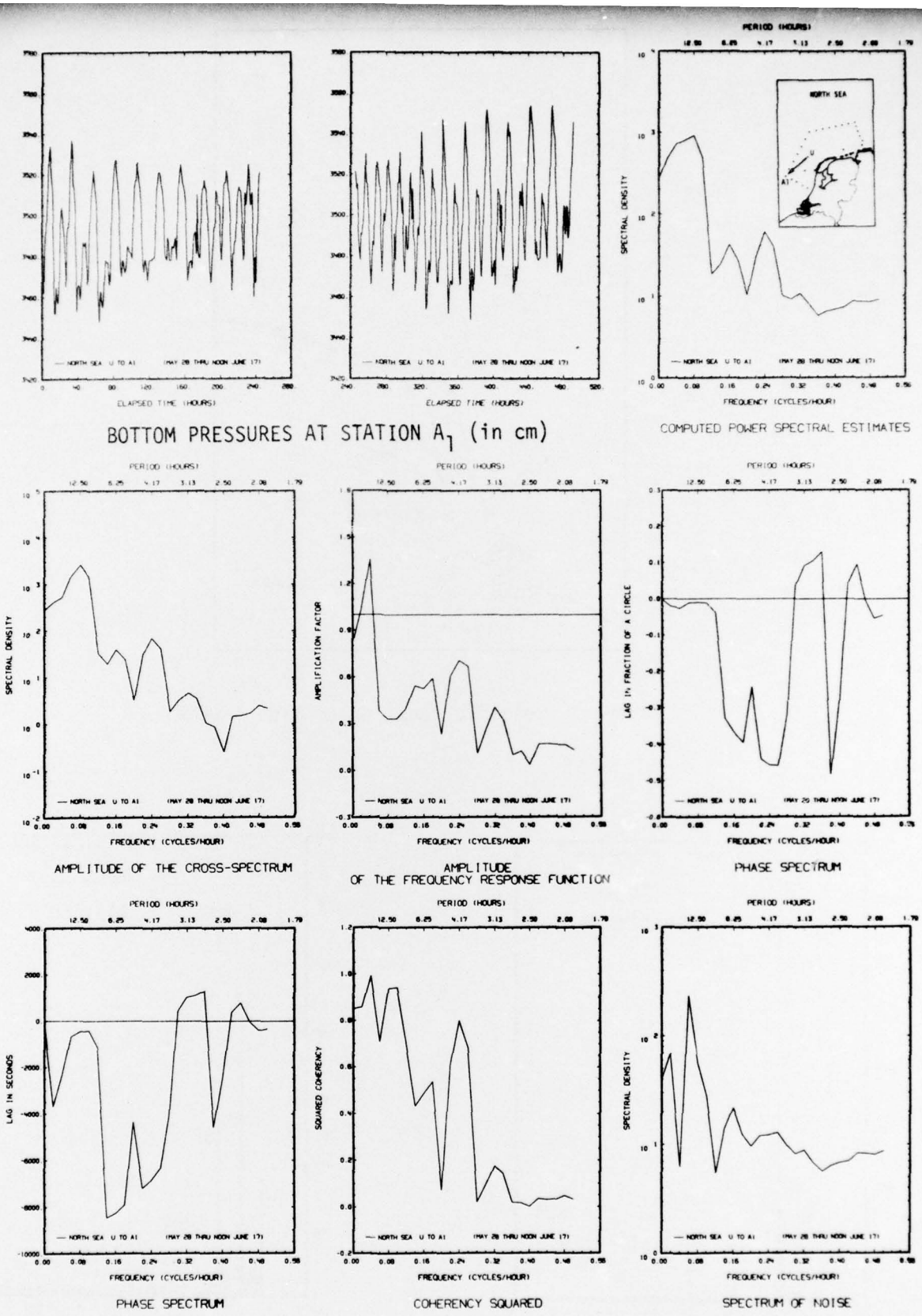


Figure 3 Frequency response analysis between Station U (reference station) and Station A1

tide (M_6 at f = .24 hr^{-1}), contain much less energy than the semi-diurnal tide at f = .08 hr^{-1}. This can also be found in the cross-spectrum between the reference station U and A_1 in the first graph in the second row of Fig. 3. The amplitude of the frequency response estimate between U and A_1 shown in the middle graph, indicates that the amplification factor is 1.36 for the diurnal tide (f = 0.04), but only 0.3 for the semi-diurnal tide (f = 0.08). The phase of the response function is shown in the third graph of the second row in fractions of a circle. The phase is also shown in the bottom row as the lag in seconds. The computed square coherency and the spectrum of the uncorrelated components are shown in the bottom row.

The spatial distribution of the amplitude and phase of the frequency response function and the spatial distribution of the coherency of the records for the semidiurnal component from the reference station U and the boundary stations are shown in Fig. 4. The decrease in amplification near station A_2 indicates the passing of the amphidromic point located approximately halfway between the English and Dutch coasts (see Proudman and Doodson, Fig. 5, Ref. 9). The amplitude, phase, and squared coherency of the computed frequency response function for the quarter-diurnal harmonic (f = 0.16/hr) along the boundary network are shown in Fig. 6.

Frequency response function for points between gauges can be interpolated for models of different grid size. The impulse response function h(k) between the reference station U and the boundary stations is obtained by inverse Fourier transform from the co-, quad-, and auto-spectra of the reference station.

$$\hat{h}(k) = \frac{\Delta f}{2} \left\{ \sum_{h=0}^{m} \left[\bar{\hat{C}}_{sy}(h) / \bar{\hat{P}}_{xx}(h) \right] \cos \frac{hk\pi}{m} + \sum_{h=0}^{m} \left[\bar{\hat{Q}}_{xy}(h) / \bar{\hat{P}}_{xx}(h) \right] \sin \frac{hk\pi}{m} \right\} \qquad (22)$$

for k = 0, 1, 2, 3 ... m
h = 0, ... m

The water levels at these boundary points for any future condition can then be generated optimally from the measured information at station U (or from any other single station) by the convolution formula:

$$\hat{y}(n\Delta t) = \Delta t \sum_{k=0}^{n} x(k\Delta t)\, \hat{h}(n\Delta t - k\Delta t) \qquad (23)$$

RECONSTRUCTION OF BOUNDARY INFORMATION

The aforementioned approach was used for reconstructing the open boundary information of a two-dimensional mathematical model of Jamaica Bay, New York City, U.S.A., as shown in Fig. 7.[2] During a large-scale field observation of water quality for comparing simulated with observed

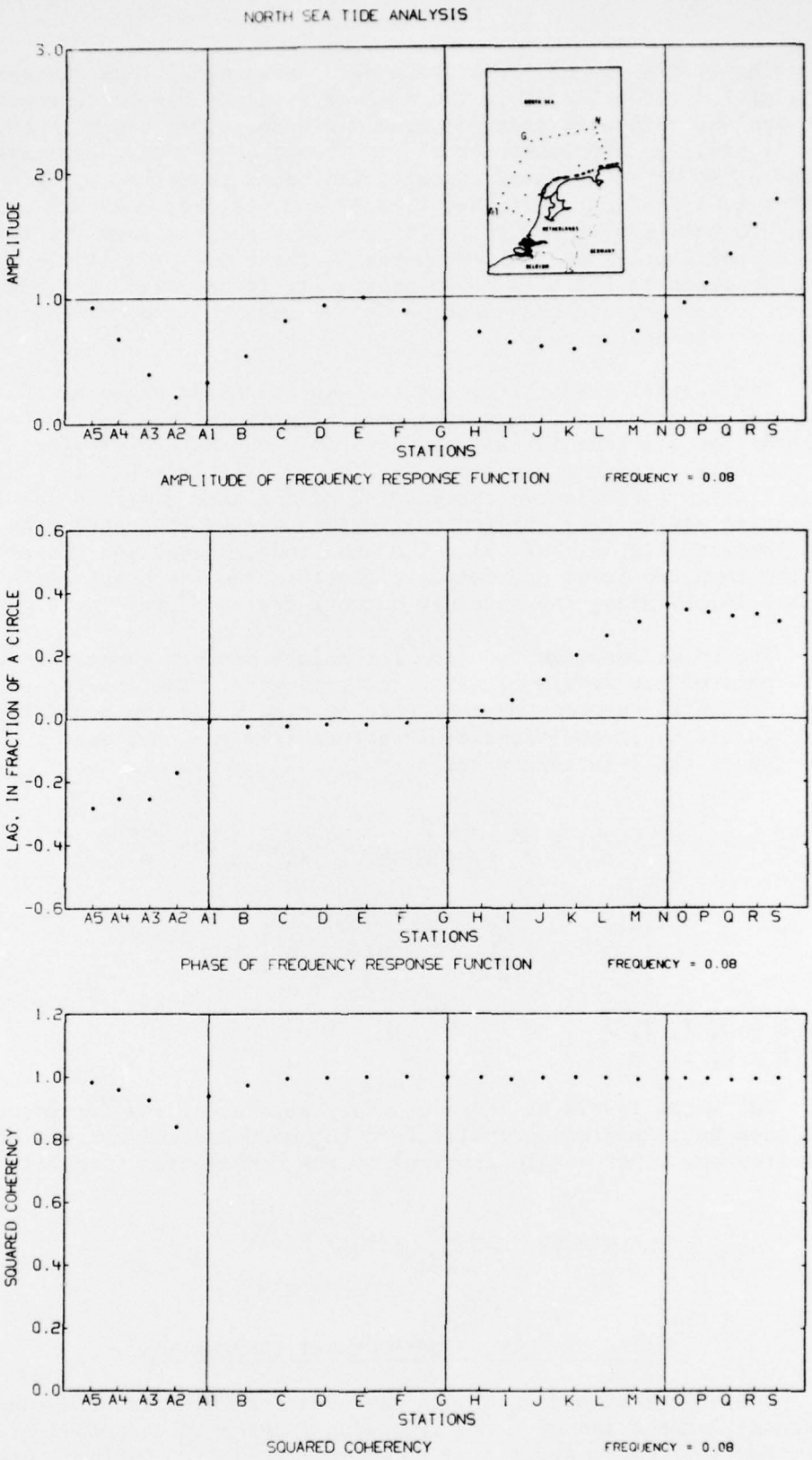


Figure 4 Amplitude, phase, and squared coherency of the computed frequency response function for the semi-diurnal tidal component (frequency = 0.08/hr)

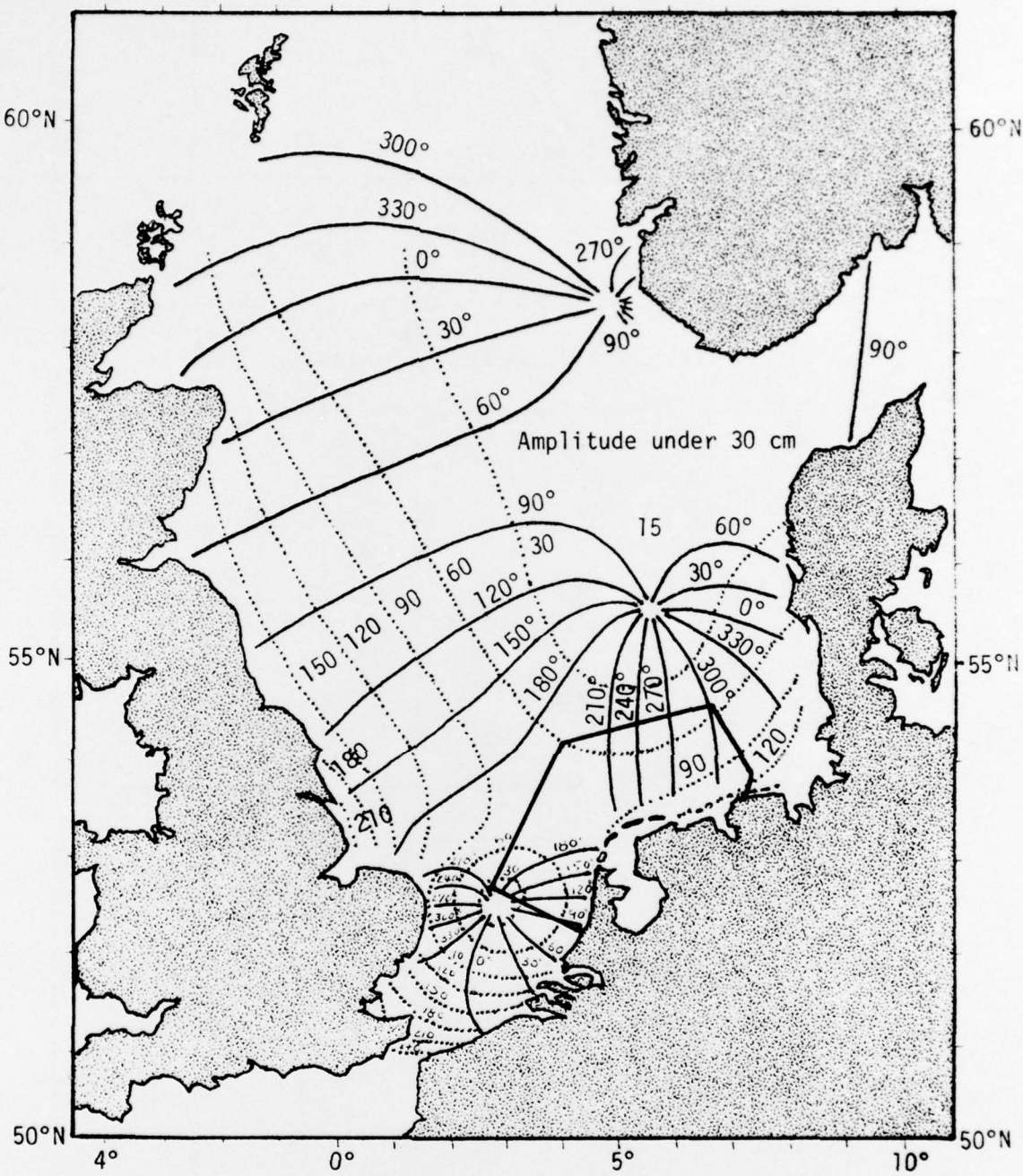


Figure 5 Cotidal times, observed amplitudes, and amphidromic points near the area of the coastal model for the M_2 component (from Proudman and Doodson, 1924)

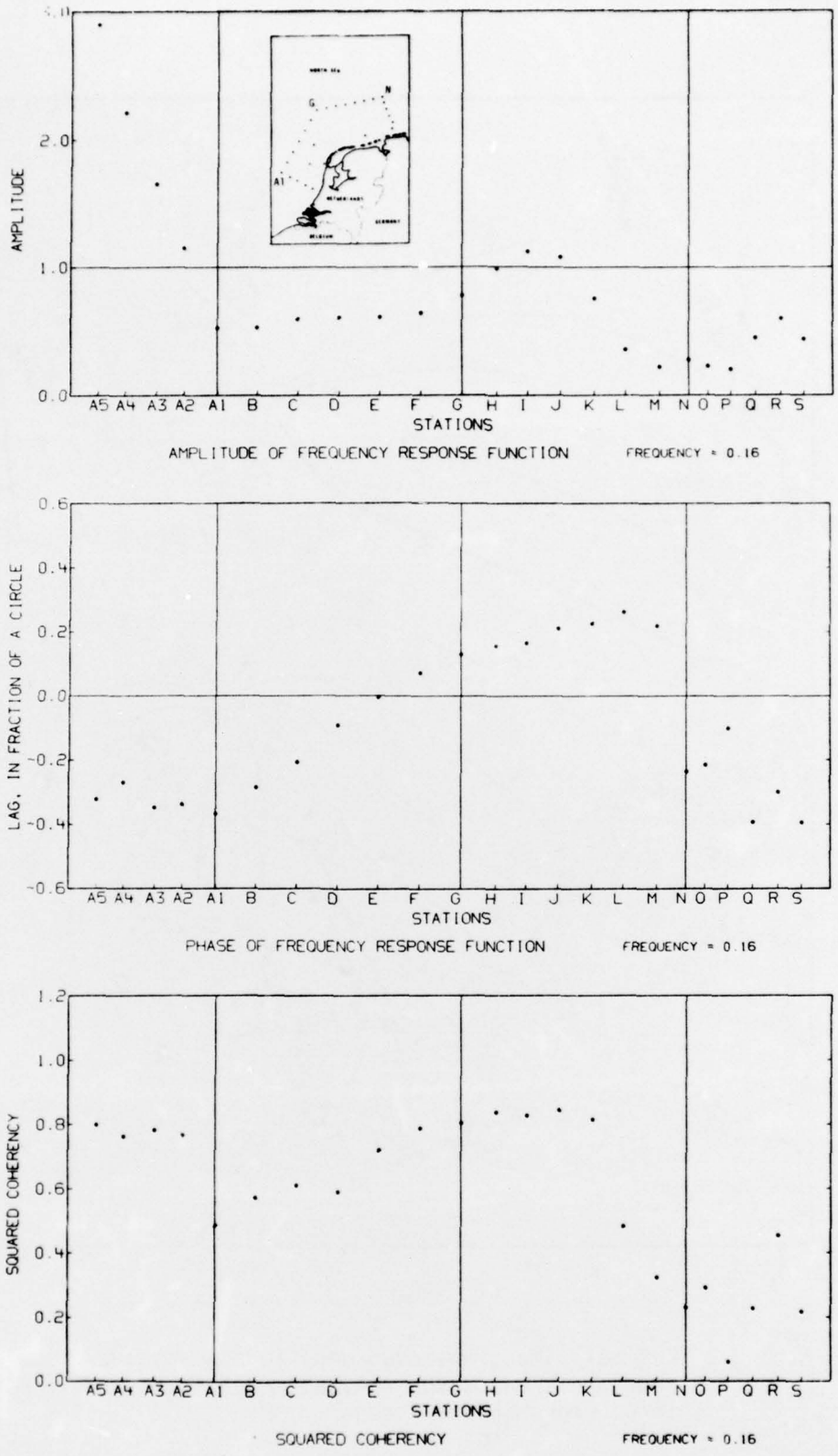


Figure 6 Amplitude, phase, and squared coherency of the computed frequency response function for the quarter-diurnal tidal component (frequency = 0.16/hr)

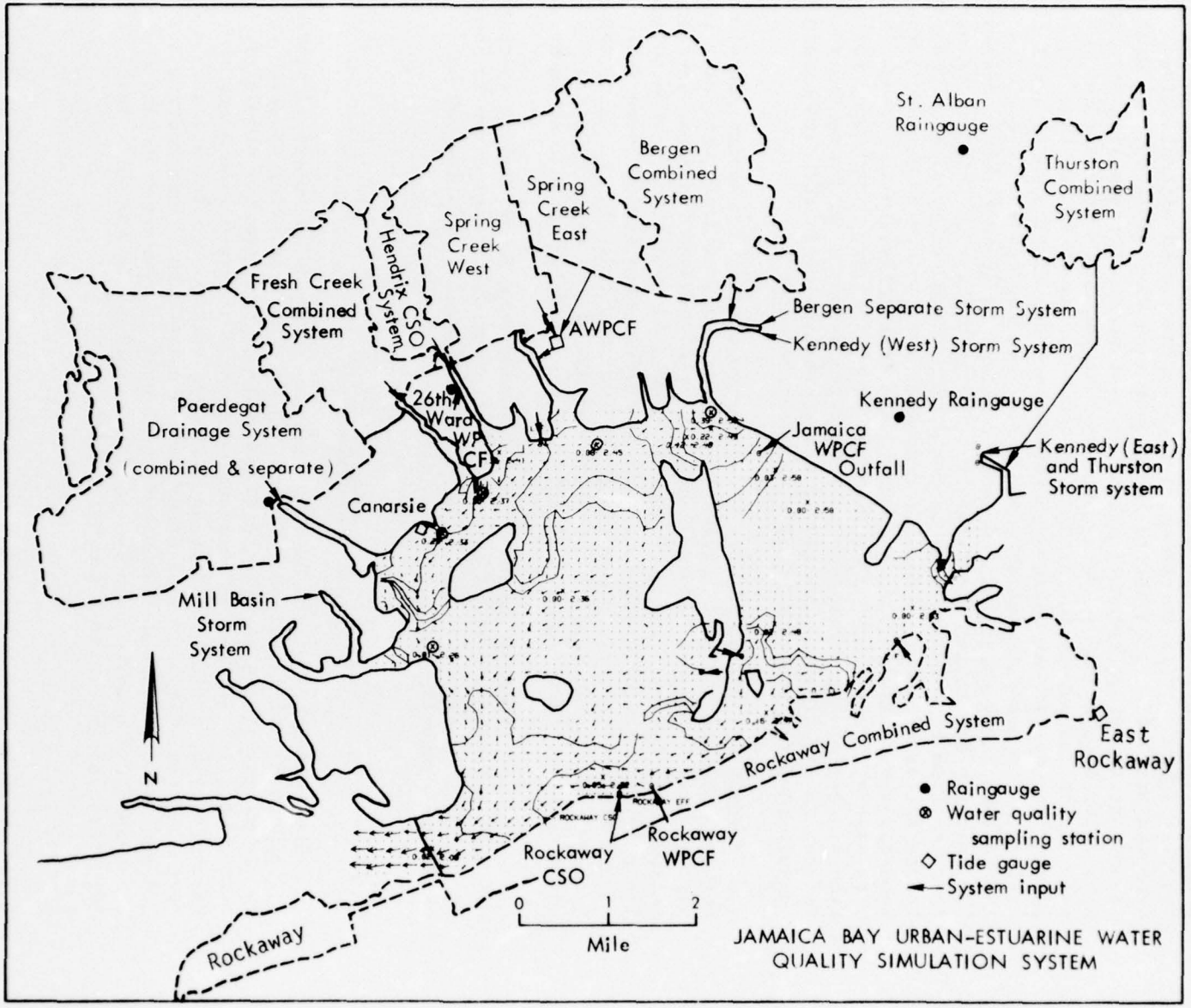


Figure 7 Jamaica Bay Urban-estuarine Water Quality Simulation System (Ref. 2)

pollutant distribution after a rainstorm, the tide gauge in the bay at Canarsie (northwestern corner of Fig. 7) malfunctioned without being detected until after the entire field operation was completed. It was found that the gauge became inoperative just before the sampling started with a few days of usable data prior to that period. Without tide information, no meaningful simulation could be made. Rather than request another survey, it was decided to reconstruct the missing water level time history. The nearest available tide data covering the entire period was the East Rockaway gauge located on Long Island (southeast corner of map). The only way for making the simulation is to derive the response (transfer) function from East Rockaway to Canarsie with the mutually available data before the experiment. Secondly, response function can be derived between Canarsie and Rockaway (open boundary) with data collected in October 1970 with high accuracy. Once these response functions are determined, the time history of the water levels at the open boundary can be reconstructed (either in the frequency domain by transformation or in the time domain by convolution) for the water quality simulation period.

Figures 8a, 8b, and 8c are the computed frequency response functions between Canarsie and East Rockaway using the group of data just prior to May 31, 1972. Figures 8d, 8e, and 8f are the response functions between Rockaway and Canarsie using October 1970 data (dotted lines). The solid lines in this set of graphs are results derived from another numerical simulation of tidal flows between these two stations. Figure 8g is the reconstructed water level (by transformation) at the model boundary using data observed from May 31 through June 3, 1972, at East Rockaway.

With the boundary information determined, the water quality simulation can then be carried out. A typical constituent distribution map is shown in Fig. 9. Detailed discussion of this particular simulation can be found in Leendertse and Liu.[2]

SUMMARY

The usefulness of the network response function in numerical simulation is threefold: (1) future conditions can be simulated using observed water levels at any single location; (2) boundary information for models of different grid size can be obtained by spatial interpolation along the boundary line; and (3) missing data at any location can be estimated optimally (in a least square error sense) using data at neighboring station and the network response functions. The uses of response function and cross-spectral density function to make numerical or hydraulic model adjustment are discussed elsewhere.[2,3]

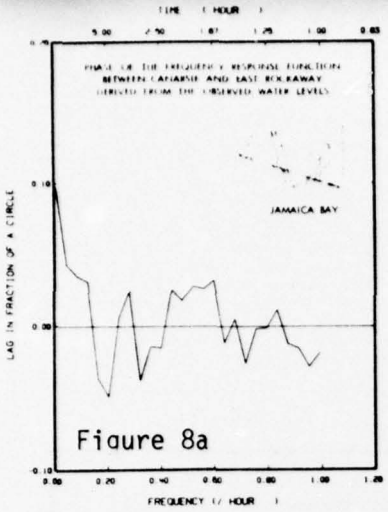


Figure 8a

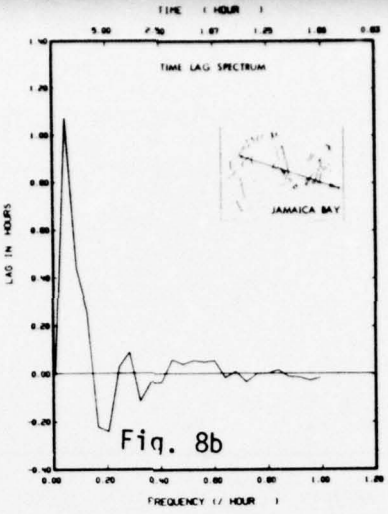


Fig. 8b

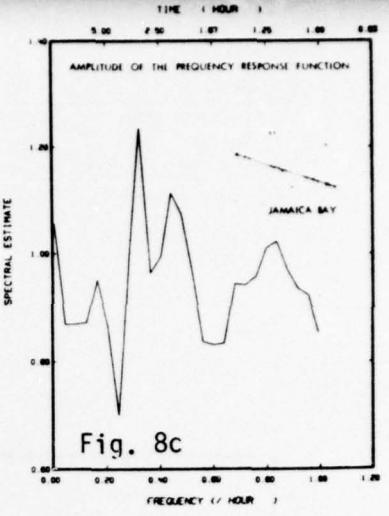


Fig. 8c

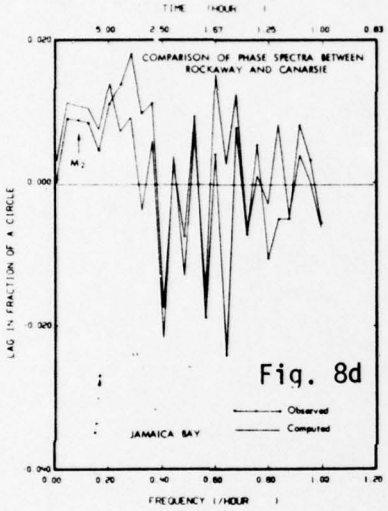


Fig. 8d

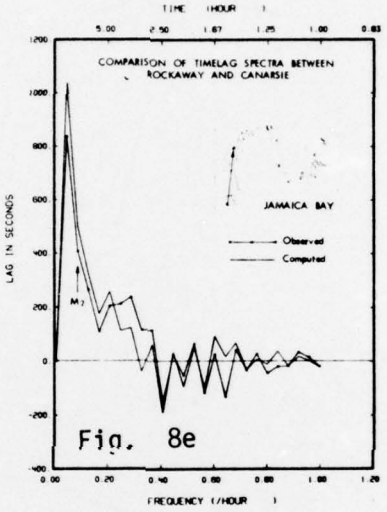


Fig. 8e

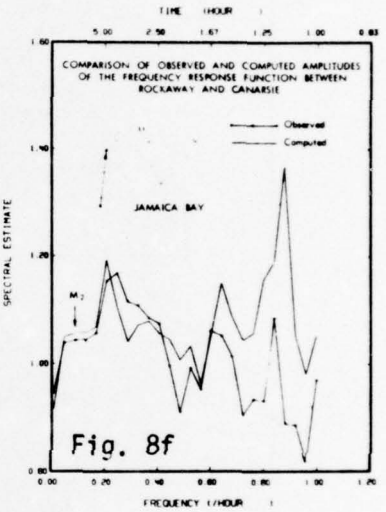


Fig. 8f

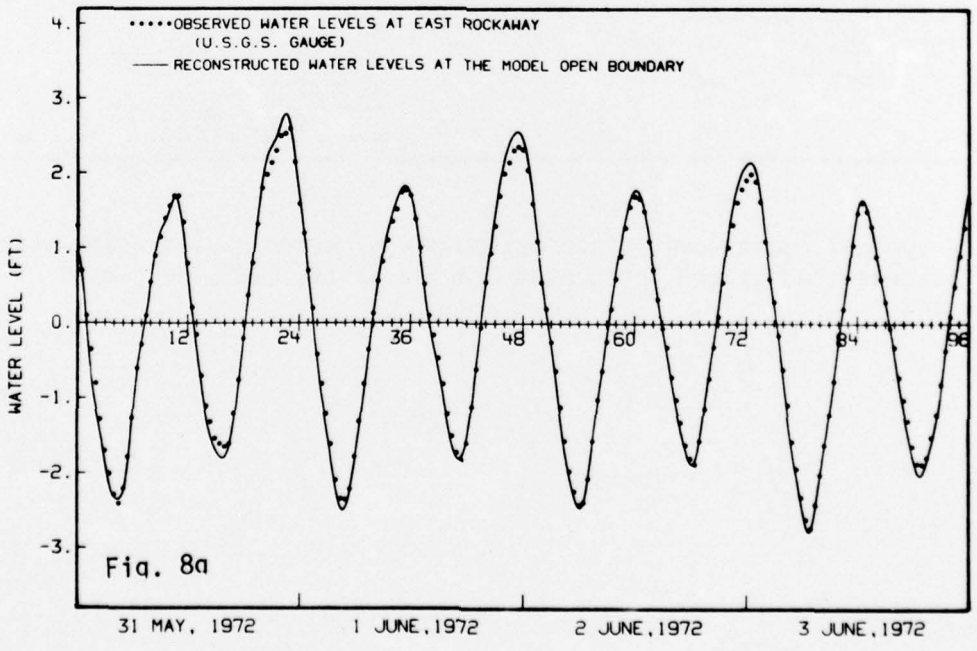


Fig. 8a

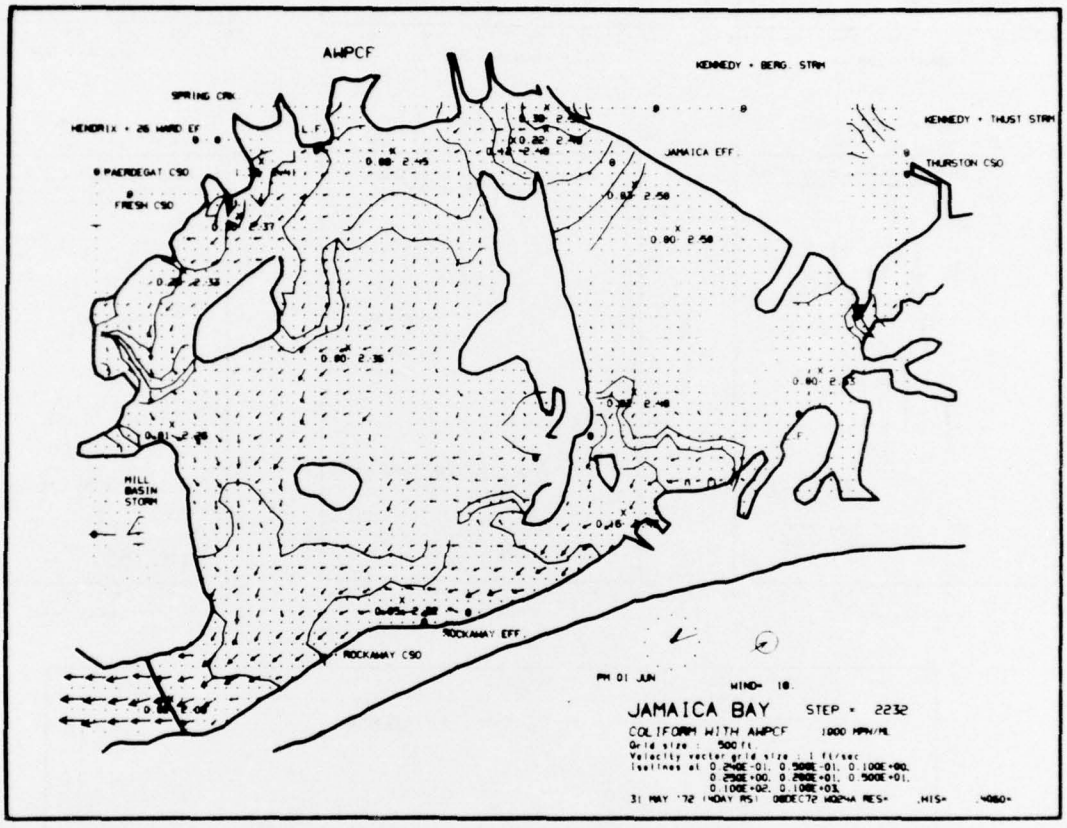


Figure 9 A typical constituent distribution map of water quality simulation as generated by the Integrated Graphic System (IGS) developed at Rand

REFERENCES

1. Liu, Shiao-Kung, *Stochastic Simulation and Control of Urban-Estuarine Water Quality Systems*, The Rand Corporation, R-1622-NYC (in preparation).

2. Leendertse, Jan J., and Shiao-Kung Liu, *A Water-Quality Simulation Model for Well Mixed Estuaries and Coastal Seas: Vol. VI, Simulation Observation, and State Estimation*, The Rand Corporation, R-1586-NYC, September 1974.

3. Leendertse, Jan J., and Shiao-Kung Liu, *Comparison of Observed Estuarine Tide Data with Hydraulic Model Data by Use of Cross-Spectral Density Functions*, The Rand Corporation, R-1612-NYC, September 1974.

4. Jenkins, G. M., and D. G. Watts, *Spectral Analysis and its Applications* Holden-Day, 1968.

5. Goodman, N. R., *On the Joint Estimation of the Spectra, Co-Spectrum and Quadrature Spectrum of a Two-Dimensional Stationary Gaussian Process*, Engineering Statistic Laboratory Scientific Paper 10, New York University, 1957.

6. Wiener, N., *The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*, John Wiley & Sons, Inc., New York, 1949.

7. Lee, Y. W., *Application of Statistical Methods to Communication Problems*, Lincoln Laboratory, Technical Report 181, 1950.

8. Tukey, J. W., "An Introduction to the Calculation of Numerical Spectrum Analysis," in *Spectral Analysis of Time Series*, Bernard Harnes, ed., John Wiley & Sons, Inc., New York, 1967, pp. 25-46.

9. Proudman, J., and A. T. Doodson, "The Principal Constituent of the Tides of the North Sea," *Phil. Trans. Roy. Soc.* A224 (London, 1924).